Retelling Narrative in 360º Videos: Implications for Audio Description

Anita Fidyka & Anna Matamala

The aim of this article is to question whether the approach for producing audio description (AD) in 2D films needs to be revisited for 360° narrative videos, a new media format characterized by its immersive capacity. To provide answers, a two-step research methodology was designed. First, an extensive literature review was performed. The data obtained during the first step was then used to design and carry out focus groups. The first part of the article discusses the findings from the literature review, comparing standard narratives with 360° narrative videos. It draws some conclusions for audio describers in relation to AD content selection, a key task in the translation of visuals into words. In the second part of the article, data obtained from the focus groups held with describers and AD users is presented. The results suggest possible approaches to AD for 360° content, such as the use of spatial sound and elements of interaction.

Keywords: audiovisual translation, media accessibility, immersive environments, CVR, user-centered methodology, audio description

1. Introduction

Virtual environments (Slater and Usoh 1993) are increasingly present in our society to inform, educate, or entertain audiences. Under this umbrella term, three main types of content can be found: Virtual Reality (VR), Augmented Reality (AR) and Mixed Reality (MR) [authors]. Within the field of VR, one can find 360º videos, which contain content based primarily on videos as apposed to computer-generated images (EBU 2017). 360° videos offer a 360-degree horizontal field of view and 180-degree vertical field of view and are becoming a new storytelling medium (Damiani 2017) for their interactive potential. Storytelling techniques in 360° videos are currently being researched by different scholars (see Section 2) and filmmakers who are experimenting with so-called Cinematic Virtual Reality (Gödde et al. 2018). 360° videos can have one linear story or they can propose navigational alternatives (Reyes, 2017, Lee 2017). In such videos, the viewer decides which viewing path to take, for instance, what decision the protagonist has to make. It is, however, the first group of 360º videos that will be the focus of this article.

To date, most of the 360º content are non-fiction, location-based programs, such as the news and current affairs (Jones 2017, 171) whereas fictional content such as animations, short films or episodic series (Dooley 2017, 164) are being developed to a lesser extent. Thanks to this media format, audiences can follow important social events from their location (EBU 2017, 9), or interact with these locations in a way that would not be possible in real life. For example, a journey to the center or orbit of our planet can be made. This is why the concept of presence, defined as "psychological sense of being there" (Biocca 1997, 18) is central to experiencing this media format. In other words, users interacting with 360º videos are expected to feel immersed in the content presented.

Taking into account the human right to access culture and information, virtual reality content, including 360º videos should be rendered accessible. This access can be provided by a

variety of different access services, which can be viewed as an instrument with which to ensure that human rights are being upheld (Greco 2016).

360º videos are multimodal (Braun 2008, 3), which means that they create meaning from the combination of and interaction between different semiotic channels: images, dialogues, sounds and on-screen text. As the visual channel is involved, they are not fully accessible for persons with sight loss. Audio description (AD) and additional access services such as audio subtitles (AST) could then be used to render this new type of audiovisual content accessible to the visually impaired target audience. In general terms, AD is a type of intersemiotic translation that transfers the visual channel to a verbal mode (Maszerowska et al. 2014; Fryer 2016). In audiovisual products, such as films, theatre, opera, sports or music events, a description of the visual components is provided along with certain sound components that would be difficult to recognize without their verbal description (Remael 2012). It is primarily created for audiences with vision loss as it enables this group to create a mental representation of the visual layer, but it can also be benefited from by other groups, e.g. foreign language learners or immigrants. In terms of narrative productions, it allows its users to recreate the original story.

As far as the production of AD in films is concerned, and similarly for any type of translation process, ADLAB guidelines (Remael et al. 2014) state that audio describers should conduct a source-text analysis before drafting a script. Such analysis is a two-step process. In the first step, the describer analyzes the source-text in order to understand the story the filmmaker intends to convey and to determine which specific narrative elements are used to do so. This allows for a second step, i.e. deciding which elements to include in the description to allow the audience to recreate the story. Guidelines currently only offer help with content selection for standard films, which is identified as one of the main issues in AD (Vercauteren 2016) but has not yet been addressed in relation to 360º narrative video content.

In order to assess whether the current practices for producing AD in 2D films have to be revisited for 360º narrative videos, a two-step research methodology was chosen. As an initial research tool, a literature review was carried out to show how 360º narrative videos change the experience of traditional narratives, and to offer initial insights into content selection in the new media format. Based on the data obtained during the literature review, focus groups were designed and held. The data obtained through these two methodological tools will help to define how AD can be integrated into this media format and help to expand knowledge on this type of intersemiotic translation, beyond traditional genres and content types.

The article begins with a brief discussion of the principles of story-creation and story-recreation, followed by an overview of existing studies on storytelling in 360° narrative videos. Section 3 provides a state of the art on narratology in AD and draws some conclusions on content selection for audio describers. Section 4 describes the methodology of the focus groups and section 5 discusses the results obtained in relation to storytelling in 360º narrative videos.

2.   From Traditional Film Narrative to Narrative in 360° Videos

Before discussing previous studies about storytelling in 360° videos, this section will briefly discuss the general principles of story-creation and story-recreation, as knowledge about these two processes is a valuable tool when conducting a source-text analysis (Vercauteren 2012, Remael et al. 2014).

Narratology is a discipline that studies all types of narrative including verbal and audiovisual texts (Bal, 1985: 5) and helps one to 'understand, analyse and evaluate narratives' (ibid.). Bal differentiates between three layers in a narrative text: the fabula, the story and the text (1985: 9). The author defines the fabula as 'a series of logically and chronologically related events caused or experienced by actors and consists of the following elements: events, actions, time and locations. These elements are later ordered into a story, which is defined as 'a fabula presented in a certain manner' (1985: 5). In other words, when creating a filmic story,

filmmakers first decide what actions the characters will perform or undergo and later, how the story will be told. At this second stage they can decide, for instance, the order of the actions (chronological or non-chronological), their frequency (the actions can be shown more than once from the perspective of different characters or they can be omitted), or their speed (actions can be shown at their normal speed, in fast-forward, or in slow-motion). At this stage, the filmmakers also decide the specific traits that the characters will have and similarly, the details of the presented places. Moreover, they decide what amount of time will be given to the various elements of the fabula (Bal 1985: 9). The third layer distinguished by Bal is 'a narrative text' (Bal 1985: 5). It is a story 'told' in a medium, i.e. converted into signs. In other words, in the last stage of story-creation, the filmmakers make a concrete decision about how the story will be presented, i.e. which film techniques are going to be used.

Another process of which audio describers should be aware of is audience story-reconstruction. In general terms, the audience is presented with a concrete narrative text and must reconstruct the abstract construct used by the filmmaker to create the story. When watching a film, audiences process and interpret information by combining actions with other information and establishing temporal relations between them. Important to the story reconstruction are mental frames (Herman 2002), which the audience creates and later updates. They initially contain basic information about the characters and settings in which the actions are to take place (e.g. '2019'). Later, as audiences constantly update their mental model of the story, more details are added (e.g. 'a rainy day in October 2019'). Only when the audiences are able to create a new frame for a given event or update an already existing frame are they able to follow the story.

While the first approach allows describers to establish the most relevant narrative elements in the content, the latter allows them to know what cues are necessary to facilitate the reconstruction of the filmic story by blind and partially-sighted audiences (Vercauteren 2012,

Remael et al. 2014). In the remaining part of this section, we will look at existing studies on storytelling in 360° videos.

Studies on storytelling in traditional 2D films are numerous and the approaches are diverse. However, research on narratives in 360º videos conducted to date is scarce and something that has only recently received scientific attention (Gödde et al. 2018). As the full spectrum of possible storytelling techniques in this media format are still being defined it is not possible to delineate them. We will however provide a brief overview of the existing studies in an attempt to compare traditional films with 360° narrative content.

The first relevant aspect is the role of the audience and their level of control. Dolan and Parets (2016) enumerate four narrative forms based on user experience within the virtual world: a user can be an observer or participant and be simultaneously active or passive. Similarly with film narratives, 360° productions fall into the passive observer category (Jones 2017). The reason for this is that the audience members are placed at the center of the events, with the option to explore the virtual world without influencing the development of the narrative. In other words, narratives in 360° media formats are pre-constructed by a film director or content creator.

A second aspect worth mentioning is how the content is watched. Dooley (2017) notes the main difference is the rectangular screen with which traditional audiovisual media is typically viewed, whilst in 360° content, the storyworld surrounds the viewer. In this regard, the process of viewing 360° videos is similar to watching content in a planetarium, the difference being that members of the audience are not seated inside a dome, but wearing head-mounted displays, and discovering the storyworld through head movements. As a result, the attention of viewers towards narratologically-relevant elements cannot be directed in the same way as can be in films (Dooley 2017; Gödde et al. 2018). A study by Rothe et al. (2017) demonstrated that it is especially difficult to guide the user's attention at the beginning of a new

scene. As viewers can explore the entire visual space, they can omit significant parts of the main narration, looking instead at other parts of the storyworld. This can result in a lack of understanding of the causality of plot (Jones 2017, Syrett et al. 2017).

To guide users, dispersing the main action across several parts of the storyworld is recommended (Jerald 2015), along with the use of auditory and visual cues, which are being researched (Streit and Victoria 2017; Rothe et al. 2017; Nielsen et al. 2016). Rothe et al. (2017) tested several traditional methods of guiding the viewers' attention (salient objects, sounds, lighted objects and movement) in order to see whether they could be tranferable to cinematic virtual reality. Although the authors consider that further investigations are necessary to find the most appropriate guiding methods, the results of this study show that moving objects or lights can guide the viewing direction within the storyworld even without any sounds. Similarly, Nielsen et al. (2016) compared diegetic cues (guiding viewers' attention implicitly by encouraging them to follow a firely with their gaze) with non-diegetic cues (controlling their body orientation) and no guidance. The results indicate that such diegetic cues can guide viewers effectively within the storyworld, whilst assuming control of the viewer's action may negatively influence presence. These two preliminary studies show that diegetic cues related to movement may be used often in 360° narrative videos to indicate where the audience should look.

As far as film language is concerned, techniques such as crossfades and fade-outs may be used more frequently in this media format than in standard films. The reason for that is because some consider them to be a better method for changing the scene than cuts (Gödde et al. 2018), as changing location rapidly can be disorienting for members of audience. Recently, however, filmmakers have begun to use cuts more often, as viewers become more accustomed to them in VR (ibid.).

As cuts are being used less frequently than in standard films, the authors note that the space in 360° content undergoes a gradual transformation around the viewer, by different elements being added or removed (ibid.). Additionally, the space can contain elements of the internal lives of the characters such as emotions, dreams or thoughts (Damiani and Southard 2017, online); by means of text, visuals or colours, which appear in the storyworld. For this reason, 360° content is thought to be an appropriate media format with which to convey subjective, impressionist first-person narratives (ibid.).

All in all, storytelling in VR is still in the process of being defined. This preliminary analysis shows that although the audience in 360° content follows a linear story, pre-constructed by the filmmaker in a similar way to standard films, the differences in narrative construction still exist, such as lack of framing and a different use of guiding cues or scene transition techniques.

3. Narratology in AD Research: from Standard Films to 360º Content

Based on narratological knowledge, models for content selection in AD were developed by Kruger (2010) and Vercauteren (2016). Kruger (2010) differentiated two approaches for rendering visual information accessible for audiences with sight loss: traditional AD and audio narration (AN), located at two ends of one continuum (2010, 233). While AD reflects the visual channel objectively, subjective audio narration reflects the narratological sense of the production. AD with elements of narration and subjectivity is placed in the middle of the spectrum.

Vercauteren (2016) adopted a narratology, both structuralist and cognitive, to develop a systematic method for audio describing filmic stories, based on their most relevant narrative constituents. His research presented a 'determine-decide' approach (see Introduction), considering what narrative elements could be included in AD and how to prioritize these narrative elements based on audience needs in order to interpret and recreate stories.

Other studies conducted in the field of AD have also adopted a narratological approach. In a study for the Pear Tree Project (2012), Kruger analyzed eye-tracking and written data to discern how to select AD content that would allow the audience to reconstruct the original story in a narratologically more meaningful way. The results of the study show that elements that occupy a lot of space on the screen do not necessarily have more relevance compared to elements that are visually insignificant. Similarly, by applying a cognitive approach, Vandaele (2012) focused on how to select the most relevant AD elements. The results of his research suggest that aside from the information that allows AD users to follow the plot, the AD script should contain information about the events that can or could have happened, to allow AD users to experience suspense, curiosity and surprise. Matamala and Remael (2015) analysed the narrative of films driven by visual effects to determine whether audio descriptions written for such films require a different approach. Thanks to the analysis of 'the cinema of spectacle', the authors conclude that as such narratives aim to dazzle their audience with effects, the emphasis in the description should be placed on the prosody of the AD script, the use of metaphors and interaction with sound. In other words, *how* to describe should be given importance in this type of filmic production.

Maszerowska (2014) studied the visual codes of film language such as light and contrast and how they should be rendered in AD. The article provides a preliminary compilation of the existing solutions for rendering lighting in AD. It concludes that more consistency is needed in the rendering of light in AD script and that knowledge about the different functions of light can prove helpful when drafting AD scripts. Taking a cognitive approach, Fresno (2014) examined the reception of characters in fiction films. The results of her investigation show that user recall is favoured by limiting the information in the descriptions and dividing it into short units for delivery at different stages.

The studies related to standard content are multiple, but it remains to be seen how the current approach for providing AD could be transferred into new 360º formats. In the following sub-sections we will focus on the main constituents of film narrative: events, spatio-temporal settings, characters and film language and strategies for rendering into AD for 2D content. This will be done so with a view to assessing what strategies could be used in 360º narrative videos to allow the audience to reconstruct their stories.

3.1. *Describing Events*

As previously mentioned, audio describers should pay attention to order, duration and frequency when describing action (Remael et al. 2014). Regarding these elements, audio describers are advised, for example, to signal any instances of events that are out of chronological order, to prioritize information or describe the visuals in advance in case of events being shown on the screen simultaneously by means of a split screen or deep space composition. Also, in case an event is shown more than once at different points during a film, describers should determine whether it is repeated unmodified or with a different status and, in turn, decide whether to include the repeated event in the script (Remael et al. 2014, Vercauteren 2016).

In 360° contents, events are rarely shown out of chronological order, meaning that the action can be described in the AD without any temporal signaling (e.g. 'a year before'). Also, actions can happen simultaneously, not by means of a split screen, but by locating them in different parts of the sphere. In this regard, as is the case with standard films, describers must decide during the text analysis, which events guarantee the coherence of the story and which complement the main narrative. The possible challenges regarding the description of events may be related to the signaling via which part of the storyworld they are presented. This information must be provided to allow persons with sight loss to have a more immersive viewing experience.

3.2. *Describing Spatio-temporal Settings*

Bal's theory of narrative space (1995) defines settings as an evolving constituent of narrative. According to Bal, when a story returns to settings that were previously introduced, they can be repeated in an unchanged way (repetition), new elements can be added (accumulation) or they can undergo a transformation. In addition to providing a background, settings can have a symbolic function.

When analysing a source text, audio describers must determine whether settings are new or known, whether they are only a background or if they have a symbolic function. If settings are already known, describers have to determine whether they are the same as before or whether they have changed. In circumstances where the settings serve only as a background, a description with basic spatio-temporal information is sufficient. Regarding settings with a symbolic function, a more detailed description should be provided. If the settings return to a place that is already known, describers have to determine whether something has changed (Vercauteren & Remael 2014).

When describing settings in 360° videos, describers will have to follow the same rules as they would when describing standard content, i.e. determine where the action takes place, in which period and the function of any given setting. The changes in settings that can occur around the viewer, i.e. addition or disappearance of visuals within the storyworld were defined by Bal as 'accumulation' (Bal 1985) and should be similar for those in standard films. Such elements should be signaled in the script. In a case where a given element reappears, describers must determine whether it is unmodified or whether its status has changed. The challenge will probably relate to the prioritization of visual information in the storyworld. As users have a 360° view, the storyworld includes more elements than standard 2D content. AD should, however, fit in between the dialogues, which may not give describers enough time to describe all relevant elements.

3.3. *Describing Characters*

The issue with describing characters has been studied by several authors (e.g. Fresno 2012; Orero 2011; Vercauteren and Orero 2013). This topic has also been tackled in the AD-Verba and ADLAB projects. ADLAB guidelines (Remael et al. 2014) distinguish between three categories of characters: namely focal, secondary and background. Each requires a different approach in AD. Focal characters are the most important to the narrative, secondary characters usually have a supporting function and background characters will be less detailed in the script. When writing an AD script, more attention is normally given to dynamic characters, i.e. characters that develop remarkably throughout the narrative, as apposed to those that develop only marginally (Vercauteren 2016, 227). Audio describers also have to decide whether the physical or mental dimension of a character is more important when describing them (ibid.). Additionally, describers need to determine whether the character being presented is new or already known (Remael et al. 2014), (ibid.), as new characters must be introduced. In this case they should then provide information about their age, appearance and characteristic traits. In the case of already known characters, describers should determine whether they have changed. Another relevant aspect worth considering is whether characters are authentic or fictional, real or unrealistic and whether they serve a symbolic function, i.e. if they represent a certain group of people, social class, profession, or stereotype.

It can be assumed that the challenges faced in regards to describing characters in 360° videos will be very similar to those for standard films. The difference being, however, that two or more characters can appear in different parts of the storyworld, as is the case with actions. Persons with sight loss may therefore find it more challenging to know where they are located without adequate guidance. Moreover, as this media format seems better suited to convey the internal lives of characters and subjective narratives, there could be need for description of emotions to be further explored, as done so by Braun (2007), Mazur (2014), and in AD

guidelines (e.g. American Council of the Blind 2009, 6). Additionally, as close-ups are not possible in this media format, identification of characters' emotions may prove more challenging.

3.4. *Describing Film Language*

The term *film language* refers to what the authors of the ADLAB guidelines (Remael et al. 2014) define as "the accepted systems, methods or conventions through which a film story comes to the audience." Under this umbrella term, four broad categories can be found: mise-en-scène, cinematography, editing and sound design. The rendering of film language in audio description has been studied by various authors (e.g. Perego 2014, Fryer and Freeman 2013, Vercauteren 2016). This topic has also been studied in the ADLAB project, with the example of the film *Inglorious Basterds* (2009). When writing AD for films, describers should determine the narrative function of the techniques used by the filmmakers (Remael et al. 2014), and may want to adopt another strategy, such as including the actual filmic techniques in the description, as done so for art-house films (Szarkowska and Wasylczyk 2014, Walczak and Fryer 2017).

Although 360° videos are in their infancy, certain aspects differ from those in standard films and may impact on descriptions. As far as cinematography is concerned, the main challenge for describers could be the lack of framing, which results in a larger amount of visual detail, more than is typically present in standard films. Similarly, as different types of shots such as close-ups are not possible, prioritizing the information presented in the storyworld may prove challenging. Regarding editing, crossfades and fade-outs, which tend to appear more often in 360° videos than standard films, should be signaled in the AD script (e.g. the screen fades to black.) Other aspects such as mise-en-scène comprise "all the elements placed in front of the camera to be photographed: the settings and props, lighting, costumes and make-up, and figure behavior" (Bordwell and Thompson 1990, 410) and will be used in a similar way to standard films but the amount of visual detail could be challenging for describers.

It can be concluded from this section that content selection in 360° videos will bear many similarities to content selection in standard films. It can also be asserted that the possible challenges when describing such content can be related to a) guiding the attention of AD users towards the relevant narrative elements, and b) prioritizing information in the storyworld. In the following sections, the methodology applied in the focus groups and the participant feedback on storytelling are discussed and serve as a necessary first step in the potential reshaping of AD in new media.

4. Focus Group Methodology

The previous sections have discussed how storytelling is changing in 360º videos and have identified how this can impact on audio descriptions for this new media format. However, when designing new solutions it is advisable to include the actual users in the making process in order to bridge the maker-user gap (Greco 2018). Based on the data collected during the literature review, focus groups were designed to involve end users. This user-centric approach has been applied at the core of the ImAc project[1], in which the implementation of access services in 360º videos is being researched alongside the development of the technology. This approach is similar to the view put forward in the field of accessible filmmaking (Udo and Fels 2010a, 2010b; Romero-Fresco 2013), stating that access services should be considered already in the content production process and not as an afterthought.

The focus groups proved of value in the early stages of the project as they enabled participants to become familiarized with 360º technology. This qualitative research method

---

[1] Immersive Accessibility (ImAc) is a 30-month project funded by the European Union's Horizon 2020 Research and Innovation Programme under the grant agreement 761974 (http://www.imac-project.eu/). The research presented in this article has been developed within the framework of this project.

allowed participants to ask additional questions about the possibilities of this technology as well as to confront their views on the implementation of access services with other participants.

The first focus group took place in Barcelona, Catalonia [authors], and the second, in Kraków, Poland. Both studies had the same structure and consisted of two tasks followed by a discussion and agreed conclusions (see procedure in section 4.2). The studies were conducted in accordance with ethical procedures approved by the Ethics Committee of ANONYMIZED. Participants were informed about the aim and context of the study and were then given consent forms to be signed. All data was confidential and the identity of the participants was made anonymous.

The focus group in Barcelona (24 November 2017) was conducted with two end users, three audio describers and one technical expert. The focus group in Kraków (28 December 2017) was conducted with three audio describers and three AD users (see participant profile in section 4.1). As the literature on qualitative research design recommends a limited number of participants for conducting focus groups (Barbour 2008, 60), the inclusion of six participants proved an optimal number for the discussion. Aside from participants, a facilitator moderated each focus group.

4.1. *Profile of Participants*

Before the discussion, participants filled in a pre-questionnaire aiming at determining their demographic profile. The pre-questionnaire contained eleven questions regarding age, sex, educational background, visual impairment, use of technologies, exposure to immersive media and access to online content. Participants were aged between 25-51 years old. Three end users defined themselves as blind. Of these three, two defined themselves as blind from birth and one, as becoming blind between 5-12 years old. The remaining two end users defined themselves as partially sighted. Almost all (11) participants had a university education and one participant reported having "further education." The responses proved that all participants were

frequent technology users. Most of the participants used mobile phones (10) and laptops (10) on a daily basis, followed by PC (5), tablets (5) and TV (5). All participants were familiar with AD. Only one participant owned a device to access VR content.

4.2. *Procedure*

The actual discussion in each of the focus groups held in Barcelona and Kraków were preceded by two tasks. Task 1 began with showing participants a 5-minute 360° video and asking one of the professional audio describers to provide a live AD. The input in the focus group in Barcelona was an episode of "Polònia", a comedy show broadcasted on public Catalan television. The input used in Kraków was an interview with a ski jumper. The reason behind choosing these materials was that both comprised of actions happening all around the sphere which allowed the users to follow the main action or to ignore it by looking at their surroundings. In Task 2, participants watched the clips with spatial sound. Asking the end users about this technology was deemed important, as it gives the possibility of placing AD in different locations within the storyworld. Thanks to said sound technology in AD, persons with sight loss will be able to localize actions and other audio elements being described.

5. Discussion of Results

In the remainder of the article, the results of the focus groups are discussed, emphasizing the aspects related to storytelling. Other relevant aspects, such as more general comments relating to the production of AD in this media format, integration of audio subtitles or interaction by voice are discussed in [authors].

First of all, professional audio describers taking part in the study in Barcelona suggested that the 360º storyworld be much larger than the standard display for 2D films and that special attention be given to setting. To this end, it was suggested that the description of the main action should be provided as a priority and that AD for the surroundings would also be needed.

Therefore, a new approach for producing AD was suggested: the sphere could be divided into four to six sections, with an AD of the surroundings provided in every section. It was later added that the main narration should be paused to listen to additional recordings.

Audio describers who took part in the study in Kraków, however, deemed this medium interactive, which means that the user can choose which parts of the storyworld to see. Therefore, they considered that persons with sight loss should also be able to watch 360° content in an interactive way. They expressed an interest in the solution proposed in the focus group in Barcelona (pausing the main narration and listening to additional AD tracks), as this solution creates the possibility of interaction.

All in all, the end users present in both focus groups agreed that an AD needs to be provided for the main action. The results for the focus group in Barcelona suggest that secondary AD tracks triggered by head movements when wearing a head-mounted display could also be implemented. This approach is interesting as it could provide a solution to the prioritization of content discussed earlier in this article. With this solution, audio describers could include more information that could later be activated at the will of the user.

Interestingly, although not strictly related to the topic of storytelling, in both focus groups, blind participants asked about the possibility of not wearing a head-mounted display. For example, two blind people in the focus group in Kraków discussed the possibility of wearing headphones with a sensor that would track their head-movements instead of a head-mounted display.

Regarding content selection, audio describers in both focus groups considered it challenging. Audio describers in the focus group in Kraków agreed that deciding which elements of the storyworld should be described would depend on the content. The responses indicate that in some content, priority should be given to conveying the intention of the director or content creator, while in others, e.g. travel simulations or city tours, describing the

surroundings should be given importance. Audio describers in the focus group in Kraków agreed that taking into account the amount of possible work, the materials for audio description should be carefully selected. One participant in this focus group added that the AD should be produced in close cooperation with the content creators or it would fall entirely under the purview of the creative team.

Regarding directing audience attention toward the narratologically-relevant elements, the responses obtained from persons with sight loss suggest that the implementation of spatial sound would be of value; it could allow them to orientate themselves and locate different elements inside the storyworld. Furthermore, the responses suggest that spatial audio could improve their immersion, which is central to experiencing 360° content.

An interesting solution was proposed by one participant in the focus group in Kraków with the suggestion of preparing visiting paths to overcome the challenge of orientating themselves within the sphere. More specifically, participants could choose a 'guide me' option which would guide them through the main action and the most important descriptions of the surroundings. The advantage of this solution is that persons with sight loss would be guided through the visual scene towards the elements that are important from a narratological point of view.

Overall, participants in both focus groups were interested in the integration of AD in 360° videos and in the possibility of activating it in an interactive way. They were mostly interested in educational or training content to be employed by museums and other institutions. They also deemed this technology more adequate for short-duration video material.

Conclusions

This article aimed at discussing whether AD requires a specific approach in 360° content. To this aim, a two-step research methodology was designed in which the bibliographical review

was followed by the focus groups, and a user-centered model of translation was adopted (Suojanen, Koskinen and Tuominen 2015).

The analysis in the first part of the article shows that the main challenges faced when describing narrative in 360° videos relate to guiding users with sight loss within the storyworld and prioritizing visual information when conducting a source text analysis.

The results obtained from the focus groups suggest possible approaches to AD in 360° content that could provide solutions to the aforementioned challenges. Such approaches could include the option of pausing the video to learn more about the storyworld by listening to additional AD tracks triggered by head movements. Another possible approach that could be combined with the one mentioned above would be the integration of spatial sound, as it can enable persons with sight loss to be guided towards the relevant narrative elements. Additionally, the results from the focus groups in terms of of content selection suggest that due to the complexity of such content, the cooperation with content creators should be sought when describing such content.

The findings discussed in this article are just a starting point in this new dynamic field. The study is limited in scope due to the early adoption phase status of this media format and that the narrative possibilities of 360° videos are not yet fully delineated. It paves the way for more complex analyses with larger samples, which would combine theoretical approaches with experimental studies. Further studies could include, but not be restricted to: measuring the effect of spatial sound in AD and audio subtitling on actual audiences, measuring user presence when faced with diverging AD strategies and testing the new production workflow with professional audio describers. As the theme caught the interest of users in both Barcelona and Kraków, elements of interactivity could be tested in future reception studies.

As the field of Audiovisual Translation Studies is being reshaped by immersive content appearing in the market, the knowledge of accessibility in such content could be integrated in

the following years in updated guidelines for audio describers in addition to courses offered by institutions for the training of future professionals. The ultimate aim is that access services – AD in this case – are produced in a way that meets quality standards, and enable all audiences to enjoy innovative media such as 360º videos.

Funding

References

American Council of the Blind. 2009. "Audio description Standards." American Council of the Blind. Accessed November 21, 2018. http://www.acb.org/adp/docs/ADP_Standards.doc

Bal, Mieke. 2009. *Narratology: Introduction to Theory of Narrative*. Toronto: University of Toronto Press.

Barbour, Rosaline. 2008. *Doing Focus Groups*. London: SAGE Publications.

Benecke, Bernd, and Elmar Dosch. 2004. *Wenn aus Bildern Worte Werden. Durch AudioDeskription zum Hörfilm.* München: Bayerischen Rundfunk.

Biocca, Frank. 1997. "The Cyborg's Dilemma: Embodiment in Virtual Environments." Journal of Computer-Mediated Communication 3 (2). doi: 10.1111/j.1083-6101.1997.tb00070.x

Bordwell, David. 1986. *Narration in the Fiction Film*. Madison: University of Wisconsin Press.

Branigan, Edward. 1992. *Narrative Comprehension and Film.* London: Routledge.

Bordwell, David, and Kristin Thompson. 1990. *Film Art: an Introduction.* New York: McGraw-Hill College.

Braun, Sabine. 2007. "Audio Description from a Discourse Perspective: a Socially Relevant Framework for Research and Training." Linguistica Antverpiensia 6: 357–369. http://epubs.surrey.ac.uk/303024/1/fulltext.pdf

Braun, Sabine. 2008. "Audio Description Research: State of Art and Beyond." Translation Studies in a New Millenium 6: 14–30. http://epubs.surrey.ac.uk/303022/1/fulltext.pdf

Chatman, Seymour. 1978. *Story and Discourse: Narrative Structure in Fiction and Film.* Ithaca: Cornell University Press.

Damiani, Jesse. 2017. "Writing for VR: the Definitive Guide to VR Storytelling." VR Scout, October 2, 2017. Accessed November 21, 2018. https://vrscout.com/news/writing-vr-definitive-guide-vr-storytelling/

De la Peña, Nonny, Peggy Weil, Joan Llobera, Elias Giannopolous, Ausiàs Pomés, Bernhard Spanlang, Doron Friedman, Maria V. Sánchez-Vives, and Mel Slater. 2010. "Immersive Journalism: Immersive Virtual Reality for the First-Person Experience of News." Presence: Teleoperators and Virtual Environments 19 (4): 291–301. doi: 10.1162/PRES_a_00005

Dolan, Devon, and Michael Parets. 2016. "Redefining the Axiom of Story: The VR and 360 Video Complex." Medium, December 20, 2015. Accessed November 21, 2018. https://medium.com/@devon.michael/redefining-the-axiom-of-story-the-vr-and-360-video-complex-bee3c20d69df

Dooley, Kath. 2017. "Storytelling with Virtual Reality in 360-Degrees: a New Screen Grammar." Studies in Australasian Cinema 11 (3): 161–171. doi: 10.1080/17503175.2017.1387357

European Broadcasting Union. 2017. "Virtual Reality. How are Public Broadcasters Using it?" European Broadcasting Union, July 20, 2017. Accessed November 21, 2018. https://www.ebu.ch/publications/virtual-reality-how-are-public-broadcasters-using-it

[authors]

Fresno, Nazaret. 2012. "Experimenting with Characters: An Empirical Approach to the Audio description of Fictional Characters." In *Audiovisual Translation and Media Accessibility at the Crossroads*. Media for All 3, edited by Aline Remael, Pilar Orero and Mary Carroll, 147–171. Amsterdam: Rodopi.

Fresno, Nazaret. 2014. "La (re)Construcción de los personajes fílmicos en la audiodescripción." PhD diss., Universitat Autònoma de Barcelona.

Fryer, Louise. 2016. *An Introduction to Audio Description. A Practical Guide*. London: Routledge.

Fryer, Louise, and Johnatan Freeman. 2013. "Cinematic Language and the Description of Film: Keeping AD Users in the Frame." Perspectives: Studies in Translatology 21 (3): 1–15. doi: 10.1080/0907676X.2012.693108

Genette, Gérard. 1980. *Narrative Discourse: an Essay in Method*. Ithaca: Cornell UP.

Genette, Gérard. 1982. *Figures of Literary Discourse*. New York: Columbia University Press.

Gödde, Michael, Frank Gabler, Dirk Siegmund, and Andreas Braun. 2018. "Cinematic Narration in VR – Rethinking Film Conventions for 360 Degrees." In *Virtual, Augmented and Mixed Reality: Applications in Health, Cultural Heritage and Industry*, edited by Chen, Jessie Y.C., 184–201. Cham: Springer International Publishing. Doi: 10.1007/978-3-319-91584-5_15

Greco, Gian Maria. 2016. "On Accessibility as a Human Right, with an Application to Media Accessibility." In *Researching Audio Description: New Approaches*, edited by Anna Matamala and Pilar Orero, 11–34. London: Palgrave.

Greco, Gian Maria. 2018. "The Nature of Accessibility Studies." Journal of Audiovisual Translation 1: 205–232.

Herman, David. 2002. "Story Logic: Problems and Possibilities of Narrative." Nebraska: University of Nebraska Press.

Herman, David. 2011. "Cognitive Narratology." In *The Living Handbook of Narratology*, edited by Peter Hühn, John Pier, Wolf Schmid and Jörg Schönert, 2. Hamburg: Hamburg University Press. Accessed November 11, 2018. http://www.lhn.uni-hamburg.de/article/cognitive-narratology-revised-version-uploaded-22-september-2013

IRT Lab. n.d. "Interactive Content." *Interactive Lab Website*. Accessed December 11, 2018. https://lab.irt.de/demos/object-based-audio/interactive/.

Jerald, Jason. 2015. *The VR Book: Human-centered Design for Virtual Reality*. San Rafael, CA: Morgan & Claypool. doi: 10.1145/2792790.2792827

Jones, Sarah. 2017. "Disrupting the Narrative: Immersive Journalism in Virtual Reality." Journal of Media Practice 18 (2–3): 171–85. doi: 10.1080/14682753.2017.1374677

Kjær, Tina, Christoffer B. Lillelund, Mie Moth-Poulsen, Niels C. Nilsson, Rolf Nordahl, and Stefania Serafin. 2017. "Can You Cut It?: An Exploration of the Effects of Editing in Cinematic Virtual Reality." In Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology. New York: ACM. doi: 10.1145/3139131.3139166

Kruger, Jan-Louis. 2010. "Audio Narration: Re-narrativising Film." Perspectives: Studies Translatology 18 (3): 232–248. doi: 10.1080/0907676X.2010.485686

Kruger, Jan-Louis. 2012. "Making Meaning in AVT: Eye Tracking and Viewer Construction of Narrative." Perspectives: Studies in Translatology 20 (1): 67–86. doi: 10.1080/0907676X.2011.632688

Maszerowska, Anna. 2014. "The Verbalizations of the Effects of Light and Contrast in Audio Description." PhD diss., Autonomous University of Barcelona.

Maszerowska, Anna, Anna Matamala, and Pilar Orero, eds. 2014. *Audio Description: New Perspectives Illustrated*. Amsterdam: John Benjamins.

Matamala, Anna, and Aline Remael. 2015. "Audio Description Reloaded. An Analysis of Visual Scenes in 2012 and Hero." Translation Studies 8 (1): 63–81.doi: 10.1080/14781700.2014.943678

Mazur, Iwona. 2014. "Gestures and Facial Expressions in Audio Description." In *New Perspectives Illustrated*, edited by Anna Matamala and Pilar Orero, 179–197. Amsterdam: Benjamins.

Meister, Jan Christoph. 2011. "Narratology." In *The Living Handbook of Narratology*, edited by Peter Hühn, John Pier, Wolf Schmid and Jörg Schönert, 25, 34, 41. Hamburg: Hamburg University Press. Accessed November 11, 2018. http://www.lhn.uni-hamburg.de/article/narratology

Men, Liang, Nick Bryan-Kinns, Amelia Shivani Hassard, and Zixiang Ma. 2017. "The impact of transitions on user experience in virtual reality." In Proceedings of the IEEE Virtual Reality (VR), 285–286. Los Angeles: IEEE. doi: 10.1109/VR.2017.7892288

Moghadam, Kasra Rahimi, and Eric D. Ragan. 2017. "Towards Understanding Scene Transition Techniques in Immersive 360 Movies and Cinematic Experiences." In Proceedings of the IEEE Virtual Reality (VR), 275–276. Los Angeles: IEEE. doi: 10.1109/VR.2017.7892333

Lee, Nicole. 2017. "Pixar's VR debut takes you inside the entrancing world of 'Coco'." *Engadget*, November 15. https://www.engadget.com/2017/11/15/pixar-coco-vr/

Nielsen, Lasse T., Matias B. Møller, Sune D. Hartmeyer, Troels Ljung, Niels C. Nilsson, Rolf Nordahl, and Stefania Serafin. 2016. "Missing the Point: an Exploration of how to Guide Users' Attention during Cinematic Virtual Reality." In Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology, 229–232. New York: ACM. doi: 10.1145/2993369.2993405

Orero, Pilar. 2011. "Audio Description for Children: Once upon a Time there was a Different Audio Description for Characters." In *Entre Texto y Receptor: Accesibilidad, Doblaje y Traducción* [Between Text and Receiver: Translation and Accessibility], edited by Elena di Giovanni, 169–184. Peter Lang: Frankfurt am Main.

Perego, Elisa. 2014. "Film Language and Tools." In *Audio description. New Perspectives Illustrated*, edited by Anna Maszerowska, Anna Matamala and Pilar Orero. Amsterdam: Benjamins, 81–101.

Remael, Aline. 2012. "For the Use of Sound: Film Sound Analysis for Audio Description: Some Key Issues." MonTI – Monographs in Translation and Interpreting 4: 255–276.

Remael Aline, Nina Reviers, and Gert Vercauteren, eds. 2014. *Pictures Painted in Words. Audio Description Guidelines*. Trieste: Edizioni Università di Trieste. Accessed November 22, 2018: http://www.adlabproject.eu/Docs/adlab%20book/index.html

Reyes, María Cecilia. 2017. "Screenwriting Framework for an Interactive Virtual Reality Film." 92–102. Paper presented at the *Immersive Learning Research Network*, Coimbra, June, 92-102. doi: 10.3217/978-3-85125-530-0-15

Romero-Fresco, Pablo. 2013. "Accessible filmmaking: Joining the dots between audiovisual translation, accessibility and filmmaking." Journal of Specialised Translation 20, 201–223. https://www.jostrans.org/issue20/art_romero.pdf

Rothe, Sylvia, Heinrich Hußmann, and Mathias Allary. 2017. "Diegetic cues for guiding the viewer in cinematic virtual reality." In Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology. New York: ACM. doi: 10.1145/3139131.3143421

Screen Australia. 2017. "Storytelling in Virtual Reality (VR): Trends and Complexities." Screen Australia, January 26, 2017. Accessed November 21, 2018. https://www.youtube.com/watch?v=U1rhfPz9f8k

Sherwood, I-Hsien. 2016. "Tech and Sympathy: Is VR the Ultimate Emotive Storytelling Tool?" Campaign, April 18, 2016. Accessed November 21, 2018. https://www.campaignlive.co.uk/article/tech-sympathy-vr-ultimate-emotive-storytelling-tool/1391407

Slater, Mel, and Martin Usoh. 1993. "Representations systems, perceptual position, and presence in immersive virtual environments." Presence: Teleoperators and Virtual Environments 2 (3): 221–233. doi: 10.1162/pres.1993.2.3.221

Streit, Martin, and Jorge Medina Victoria. 2017. "Guiding Viewers' Attention in 360 Video and VR through Audio." Paper presented at the 4th International Conference on Spatial Audio. September, 9.

Syrett, Hannah, Licia Calvi, and Marnix van Gisbergen. 2017. "The Oculus Rift Film Experience: A Case Study on Understanding Films in a Head Mounted Display." In *Intelligent Technologies for Interactive Entertainment*, edited by Ronald Poppe, John-Jules Meyer, Remco Veltkamp and Mehdi Dastani, 197–208. Springer. doi: 10.1007/978-3-319-49616-0_19

Szarkowska, Agnieszka, and Wasylczyk Piotr. 2014. "Audiodeskrypcja Autorska. " [Auteur Description] Przekładaniec 28, 48–62. doi: 10.4467/16891864PC.14.004.1711

Udo, John-Patrick, and Deborah Fels. 2010a. "The Rogue Poster-Children of Universal Design: Closed Captioning and Audio Description." Journal of Engineering Design 21 (2–3): 207–221. doi: 10.1080/09544820903310691

Udo, John-Patrick, and Deborah Fels. 2010b. "Universal Design on Stage: Live Audio Description for Theatrical Performances." Perspectives: Studies in Translatology 18 (3): 189–203. doi: 10.1080/0907676X.2010.485683

Walczak, Agnieszka, and Louise Fryer. "Creative Description: The Impact of Audio Description Style on Presence in Visually Impaired Audiences." British Journal of Visual Impairment 35, no. 1: 6–17. doi:10.1177/0264619616661603.

Vandaele, Jeroen. 2012. "What Meets the Eye. Cognitive Narratology for Audio Description." Perspectives: Studies in Translatology 20 (1): 87–102. doi: 10.1080/0907676X.2011.632683

Vercauteren, Gert. 2012. "A Narratological Approach to Content Selection in Audio Description. A Strategy for the Description of Narratological Time." MonTI. Monografías de Traducción e Interpretación 4: 207–231.

Vercauteren, Gert. 2016. "A Narratological Approach to Content Selection in Audio Description." PhD diss., University of Antwerp.

Vercauteren, Gert, and Pilar Orero. 2013. "Describing Facial Expressions: Much more than Meets the Eye." Quaderns. Revista de Traducció 20: 187–199.

Vercauteren, Gert, and Aline Remael. 2014. "Spatio-temporal settings." In *Perspectives Illustrated*, edited by Anna Matamala, 61–80. Amsterdam: Benjamins.

Word count: 7620